

# SCIENTIFIC DATA

## OPEN Data Descriptor: Twelve thousand recent patellogastropods from a northeastern Pacific latitudinal gradient

Received: 13 June 2017  
 Accepted: 17 October 2017  
 Published: 9 January 2018

Sara S. Kahanamoku<sup>1,2</sup>, Pincelli M. Hull<sup>1</sup>, David R. Lindberg<sup>2</sup>, Allison Y. Hsiang<sup>1,3</sup>, Erica C. Clites<sup>4</sup> & Seth Finnegan<sup>2</sup>

Body size distributions can vary widely among communities, with important implications for ecological dynamics, energetics, and evolutionary history. Here we present a dataset of body size and shape for 12,035 extant Patellogastropoda (true limpet) specimens from the collections of the University of California Museum of Paleontology, compiled using a novel high-throughput morphometric imaging method. These specimens were collected over the past 150 years at 355 localities along a latitudinal gradient ranging from Alaska to Baja California, Mexico and are presented here with individual images, 2D outline coordinates, and 2D measurements of body size and shape. This dataset provides a resource for assemblage-scale macroecological questions and documents the size and diversity of recent patellogastropods in the northeastern Pacific.

<b>Design Type(s)</b>	observation design • image analysis objective
<b>Measurement Type(s)</b>	morphology
<b>Technology Type(s)</b>	digital camera
<b>Factor Type(s)</b>	geographic location
<b>Sample Characteristic(s)</b>	Patellogastropoda • State of California • State of Baja California • State of Washington • Mexico • State of Alaska • State of Oregon • Province of British Columbia

<sup>1</sup>Yale University, Department of Geology & Geophysics, New Haven, CT 06511, USA. <sup>2</sup>University of California, Department of Integrative Biology and Museum of Paleontology, Berkeley, CA 94720, USA. <sup>3</sup>Swedish Museum of Natural History, Department of Bioinformatics and Genetics, Stockholm 10405, Sweden. <sup>4</sup>University of California Museum of Paleontology, Berkeley, CA 94720, USA. Correspondence and requests for materials should be addressed to S.S.K. (email: sara.kahanamoku@berkeley.edu).

## Background & Summary

Body size is one of the most important aspects of organismal form and function, and influences a broad array of physiological, ecological, and evolutionary processes<sup>1</sup>. Environmental controls on body size have been studied in many different groups for over a century<sup>2,3</sup>, particularly along latitudinal gradients, which serve as natural experiments in which to study the effects of abiotic change on organismal and assemblage size distributions<sup>4,5</sup>. There is a large amount of variance in body size distributions at all scales, which may be difficult to examine using traditional methods, many of which are time-intensive. Automated methods of measuring body size, developed more recently, have the potential to incorporate individual data from large numbers of specimens. Speeding up morphological measurements helps to facilitate the collection of large datasets that capture the full range of variation within and among communities, which may then be used to elucidate the factors that determine body size distributions across space and time<sup>6</sup>. Here we present a comprehensive database of the size and shape of Patellogastropoda (true limpets) from Baja California, Mexico to Alaska. Patellogastropods are a globally distributed group, comprising ~400 species<sup>7</sup>, that are nearly ubiquitous in intertidal habitats. They thus provide an interesting comparison to fully marine groups such as bivalves that have been the focus of previous analyses of marine body size distributions example refs 3,6. Further, because they may be affected by changes in both atmospheric temperature and the temperature and pH of ocean waters<sup>8–10</sup>, establishing baseline gradients in patellogastropod body size can help to detect and diagnose the causes of future changes in the spatial structure of body size distributions<sup>11,12</sup>.

Breakthroughs in imaging and digitization<sup>13</sup> have made the rapid collection of individual body size data feasible. A recent study of assemblage-level microfossil size and shape presented and utilized technological advances in imaging and novel open-source software, *AutoMorph*, to generate specimen-level data including individual images, 3D hulls, and 2D and 3D morphometric data using photogrammetry<sup>14</sup>. In this study, we used a modified version of *AutoMorph* that includes image processing capacities for macroscopic photographic images in order to document and measure modern Patellogastropoda from Baja California, Mexico to Alaska in the University of California Museum of Paleontology (UCMP) collections. The resulting images are tagged with metadata and associated with specimen-level morphometric measurements. All morphological information, images and metadata are openly available through Zenodo (Data Citation 1). The physical specimens associated with these images are deposited at the UCMP.

Collection localities ranged from 60°06'11N (Cape Yakataga, Alaska, United States) to 22°89'56N (Cabo San Lucas, Baja California, Mexico) (Fig. 1 and Table 1 (available online only)). The northeastern Pacific was chosen because the UCMP has extensive patellogastropod collections from this region and because mean annual temperatures, seasonality, and productivity exhibit strong variation over ~5,000 km of coastline<sup>15</sup>. 1,410 lots of recent northeastern Pacific patellogastropod specimens from the UCMP collections were imaged in the summer of 2015 by S.S.K. Of these, 1,256 lots were successfully processed via *AutoMorph* for the morphometric parameters reported here. Most processing failures for the 155 unsuccessful lots resulted from inadequate focusing of stacked images during the *AutoMorph* step (*focus*) that required use of a third-party stacking software. Each lot contained between 1 and 223 individuals (Table 1 (available online only)), and were collected between 1860–2014 by numerous collectors for a variety of reasons (e.g., taxonomic, distributional, ecological).

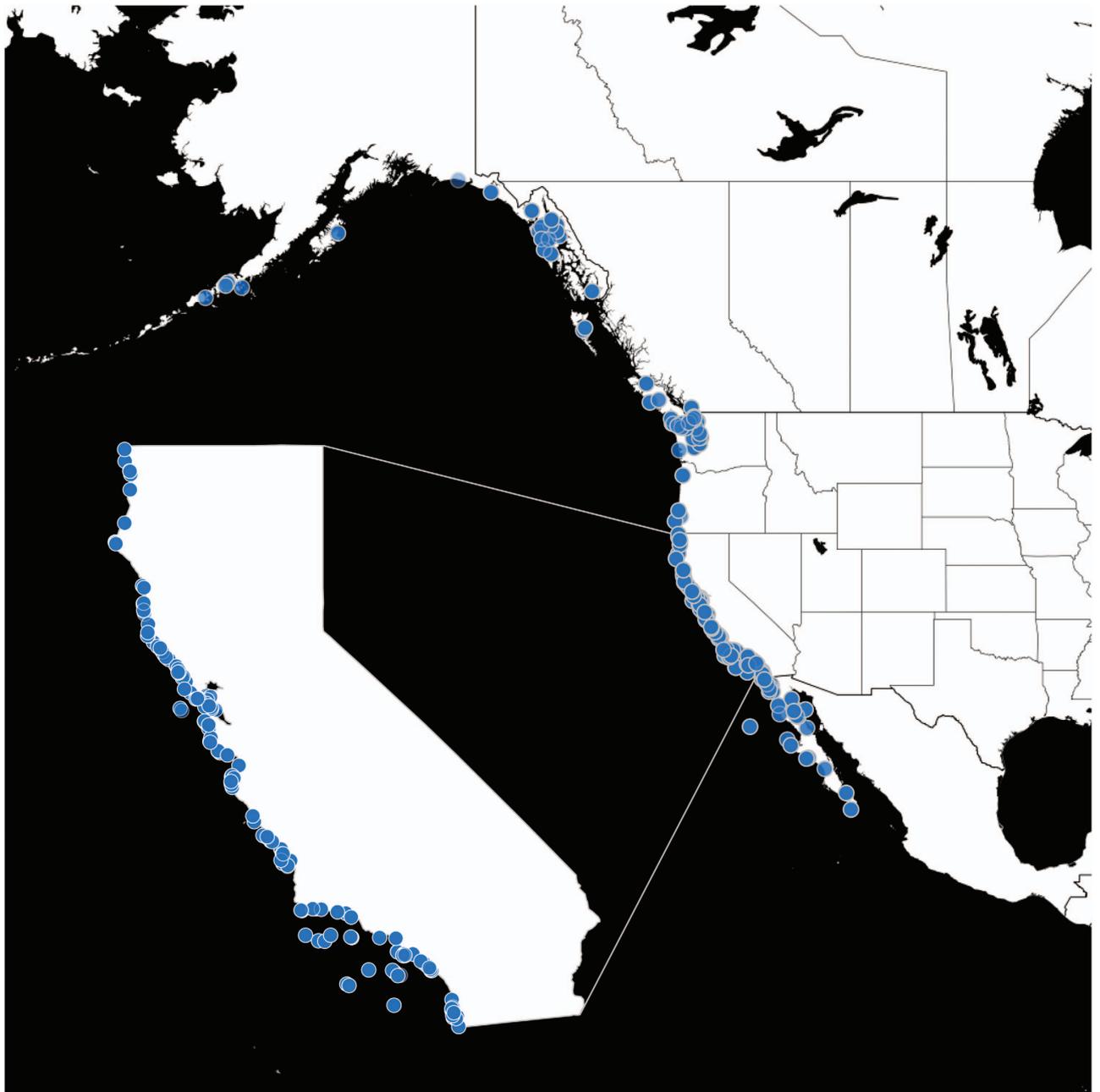
We present images and morphometric data for 12,035 individual patellogastropods, representing 30 species from 355 northeastern Pacific sites (Tables 2 and 3 (available online only)). This is, to our knowledge, the largest assemblage-scale morphometric dataset ever produced for mollusks, and documents northeastern Pacific patellogastropod size and diversity over the past 150 years.

## Methods

### Sampling

More than 90% of recent northeastern Pacific patellogastropod collections from the UCMP were photographed for this study. Because specimens in the collection had been identified to the species level and grouped by locality during accession to the museum, each lot reported in Table 1 (available online only) represents a single collection event of a single species at a single site. For this study, 1,256 of these lots from 355 unique localities between 60°06'11N and 22°89'56N (Table 2 (available online only)) were photographed and measured via digital image processing. A large number of these lots (699) were uncurated, and were assigned UCMP lot numbers prior to incorporation in the study. This preliminary curation resulted in the digitization of previously unavailable collection information (Table 1 (available online only)). Locality information was recorded as both UCMP locality numbers (e.g., 'D8919') and locality strings (e.g., 'Monterey, CA'). It should be noted that multiple UCMP locality numbers may be assigned to a single site, as samples are assigned numbers at the time of accession that reflect both the locality as well as the accession group. Locality strings were matched to latitudinal and longitudinal coordinates using GEOLocate<sup>16</sup> for all but 40 lots, and coordinates were rounded to the nearest 0.1 degree to account for potential measurement uncertainty (Table 1 (available online only)).

The UCMP recent mollusk collections date from the Geological Survey of California (1860–1874) and were deposited by the State Legislature with the university in 1873 (ref. 17). J.G. Cooper was especially active in the building of the recent mollusk components of these collections<sup>18</sup>. The early survey collections were further augmented by late 19th and early 20th century collections from Henry Hemphill

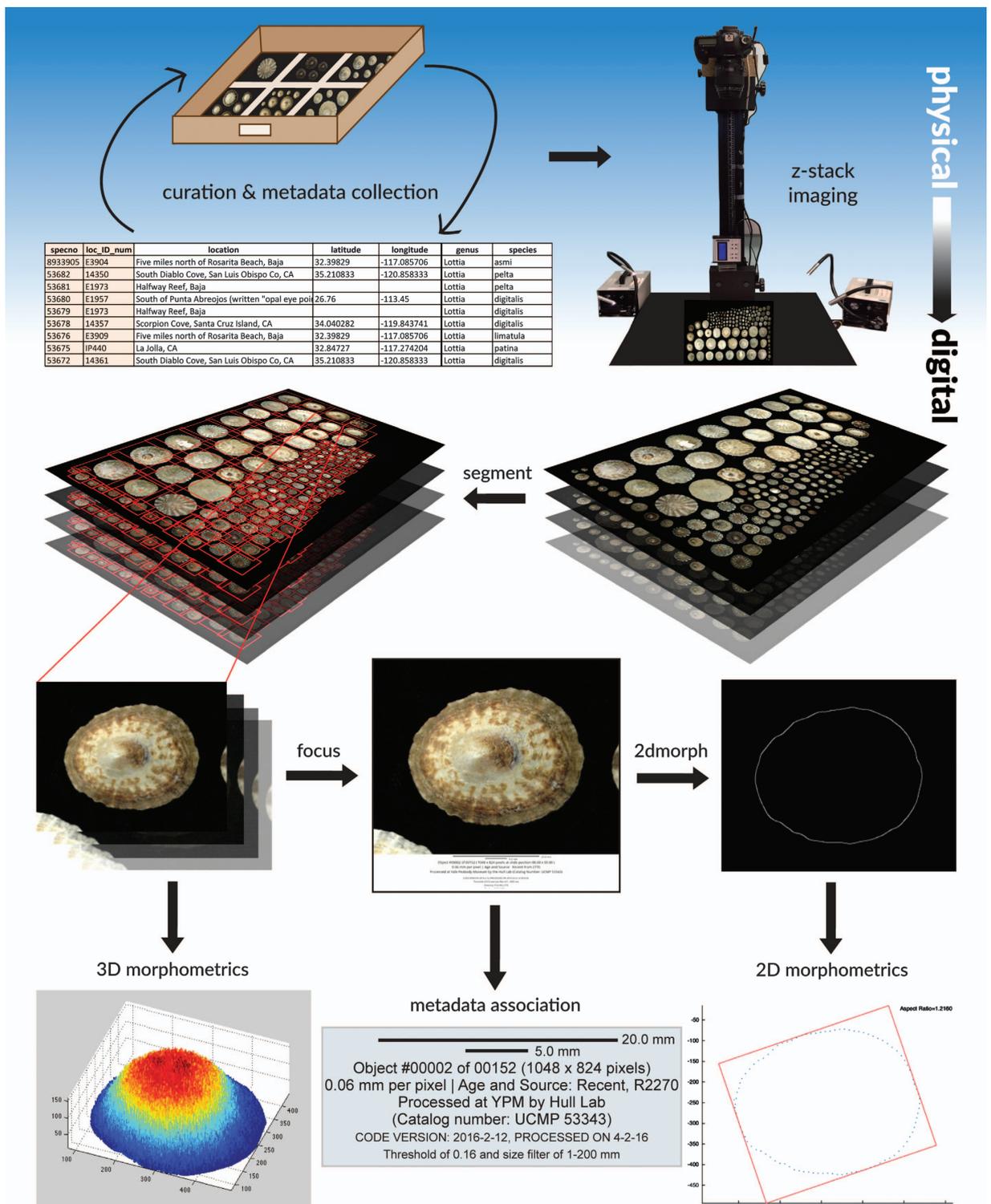


**Figure 1.** Map of northeastern Pacific with sample sites. A map of the area of the northeastern Pacific referenced in this study, with sample localities in blue. California sites are further highlighted, as the majority of localities (216 of 355) from which specimens were collected were within California.

(1830–1914)<sup>19</sup> and Josiah Keep (1849–1911)<sup>20</sup>. Additional large northeastern Pacific collections were added in the mid 20th century with the acquisition of the Eugene Coan and Rudolph Stohler mollusk collections from the University of California Davis<sup>21</sup>. In addition to general mollusk collections, collections focused exclusively on patellogastropods (Gulf of California through Alaska) were made as part of faculty and graduate research projects, including those of Avery Ransome Grant Test<sup>22</sup> (1933–1945), Henry K. Fritchman<sup>23</sup> (1950s), and David R. Lindberg and students (1982–2014). These collection activities over the last 150+ years produced the systematic collection of northeastern Pacific patellogastropods reported on here.

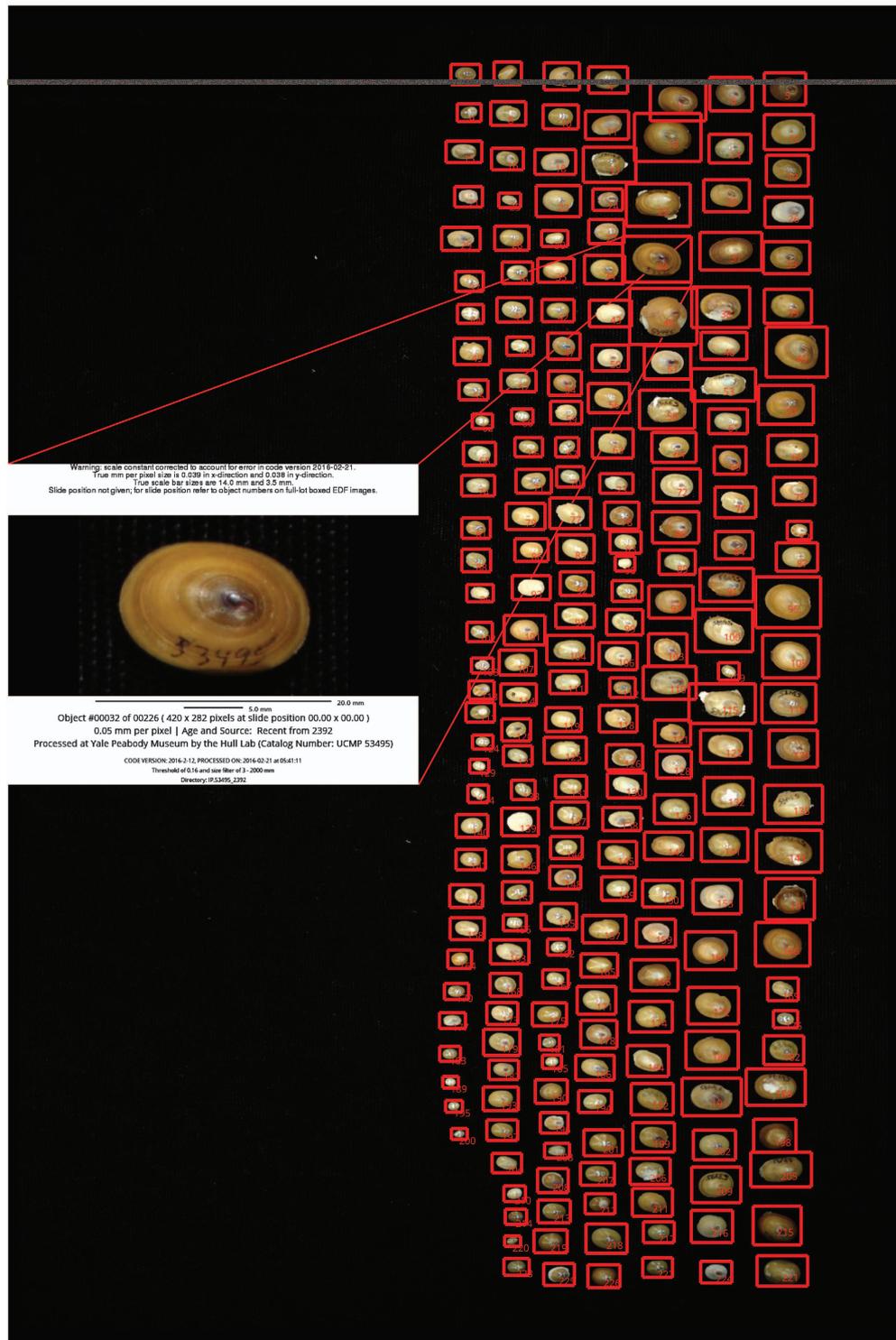
#### High-throughput imaging

Images were taken using a Canon EOS 5D Mark III camera and a Canon EF 100 mm f/2.8 Macro USM lens. Camera settings were manually optimized for macro imaging: an aperture of 2.8 and ISO of 200 were used for all images. Image settings and file storage were remotely controlled by a laptop computer



**Figure 2. Digitization workflow overview.** All physical specimens were curated and their metadata catalogued prior to imaging. Specimen lots were imaged in Z-stacks (Row 1). Following digitization, individual objects were identified from these lot images, and each object was cut out of the Z-stack and labeled with metadata using the *segment* module of *AutoMorph* (Row 2). These Z-stack images were then passed through the *focus* module to create 2D EDF images for each individual object, and metadata labels were again associated with each individual's 2D EDF (Row 3). An enlarged example of this image-associated metadata is provided (Row 4, center). Following the *focus* step, 2D EDFs were passed through the *run2dmorph* module to generate 2D morphometric data (Rows 3 and 4) while object Z-stacks were retained for future 3D morphometric work (Row 4).

Warning: scale constant corrected to account for error in code version 2016-02-21.  
 True mm per pixel size is 0.039 in x-direction and 0.038 in y-direction.  
 True scale bar sizes are 14.0 mm and 3.5 mm.  
 Slide position not given; for slide position refer to object numbers on full-lot boxed EDF images.



**Figure 3. Example labeled image.** An example overview (background) and individual image (inset). Objects are boxed and labeled with their respective object numbers, which are then used to identify individual images. Both overview and individual images are tagged with associated metadata. The images described here contain additional warning labels that provide users with corrected metadata (i.e., scaling constant and, when applicable, locality information). Shown here: UCMP 53,495 (*Lottia insessa*,  $N=223$ ), with the EDF of object #32 inset.

using the Canon EOS Utility 3 program. The camera was mounted on a Cognysis Inc. Stackshot Automated Focus Stacking Rail, which in turn was mounted on a camera stand (Fig. 2). StackShot was used to standardize the distance between image planes. The Z-stack images generated by StackShot allowed for the creation of 2D extended-depth-of-focus images and 3D photogrammetry<sup>13</sup> (for use in future studies). The distance between each z-plane was set to 1 mm per step. The number of planes varied according to the maximum height of the tallest specimen in each sample. Samples were illuminated during imaging using standard gooseneck illuminators (Fig. 2). Use of these illuminators resulted in color alteration, which was corrected during imaging using the Canon EOS Utility software.

High-throughput imaging techniques, modelled after previous studies on foraminifera<sup>13,14</sup>, were utilized to maximize the efficiency of digitization. Samples were imaged as sample lots as opposed to individual specimens, with a maximum of 150 individuals per round of photography. Specimens were laid out on a uniformly dark background such that no individuals were touching. Prior to each round of digitization, scale bars were photographed in X and Y directions in order to calibrate the image scale in mm/pixel for post-processing.

### **AutoMorph: Automated morphometric post-processing**

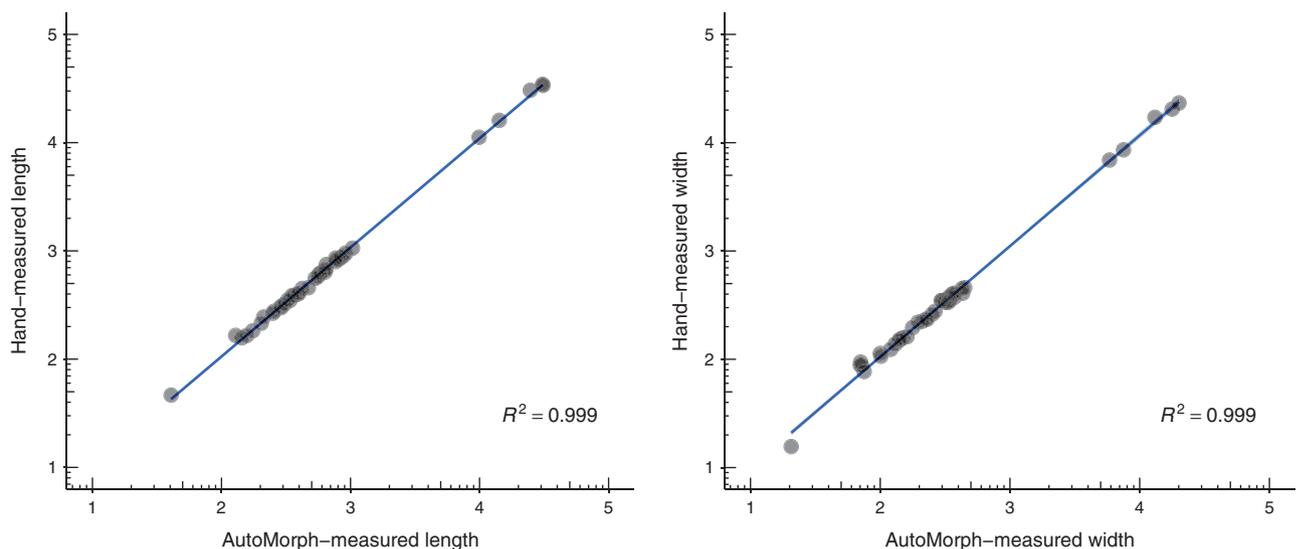
We utilized the Yale Grace high-performance computing cluster to expedite image post-processing with the *AutoMorph* software suite (available on GitHub; see *Code Availability*). *AutoMorph*<sup>14</sup> is a software package that utilizes photographic image stacks of multiple objects to extract individual object images and associated morphometric measurements and coordinates. Some aspects of the *AutoMorph* software are specific to macroscopic images and we refer to them here as the macro-version *AutoMorph*.

**Image preparation.** The macro-version *AutoMorph* contains a software module, *prepare*, which ensures that any frame shifting associated with camera zooming in the Z-dimension is eliminated. Using *prepare*, all image stacks were rotated and scaled prior to processing with *segment*.

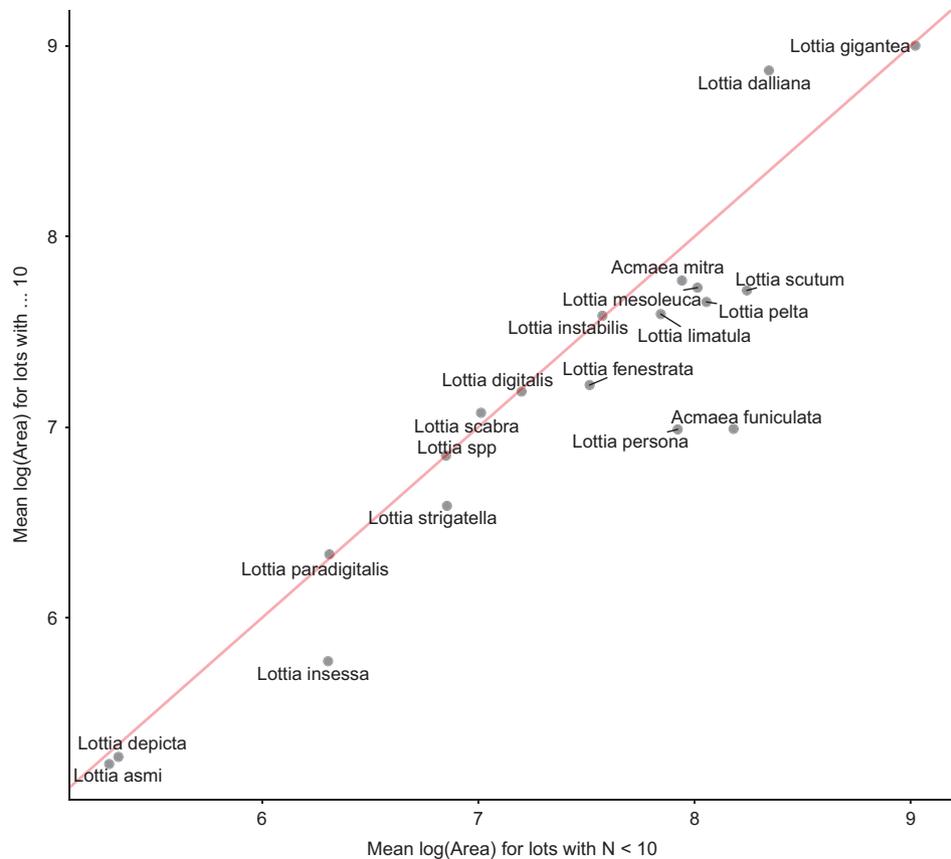
**Segmenting.** Image stacks of individual patellogastropod limpets were extracted from full-lot images using the *AutoMorph* module *segment*<sup>14</sup>. *segment* takes advantage of the color contrast between objects and the image background to identify and box each individual. These objects are then cut out of the full-lot image stack and placed into individual directories containing object-specific z-sliced images. Object numbers are assigned during processing and are labeled on both the full lot image and the individual stacks. User-defined metadata is also attached to each image by *segment*, including scale bars and processing date and location (Fig. 3).

**Focusing.** Image stacks were processed into a best 2D image, known as an extended-depth-of-focus (EDF) image, using the *AutoMorph* package *focus*<sup>14</sup>. *focus* facilitates batch processing in *Zerene Stacker* (<http://zerenesystems.com/cms/stacker>), a third-party focus stacking software, or ImageJ, an open source image processing toolkit. *focus* retains individual image z-stacks for downstream use in 3D data extraction using *run3dmorph*<sup>13,14</sup>, another *AutoMorph* module.

**2D morphometrics.** EDF images for each object were then processed to extract object outlines and 2D morphometric measurements using the *AutoMorph* package *run2dmorph*<sup>14</sup>. This 2D morphometric



**Figure 4. Size validation.** A comparison of *AutoMorph*- and hand-measured length and width (data in Table 4 (available online only)). Length  $R^2$ : 0.99; width  $R^2$ : 0.99;  $n = 41$ .



**Figure 5. Collection biases in lots with small sample sizes.** Lots with small sample sizes (defined here as lots with less than 10 individuals) may make collector biases towards larger body size more apparent. Species mean body size for small lots ( $N < 10$ ) is compared with mean body size for large lots ( $N > 10$ ); while some species means lie along the line of unity, the majority (12 of 19) appear to have larger mean body sizes when lot sample sizes are small.

software passes images through smoothing, RGB, greyscale, and black-and-white filters to extract object outlines,  $x,y$ -coordinates, and individual measurements of size and shape, such as major axis length, minor axis length, area, perimeter, rugosity, and aspect ratio. These measurements are saved in sample-specific CSV files, and object outlines are plotted on their respective images to allow for visual quality-control checks of 2D data extraction.

### Code availability

Both micro- and macro-specific versions of the *AutoMorph* software suite are freely accessible on GitHub (<https://www.github.com/HullLab/AutoMorph>). This study used *AutoMorph* v2016-02, the first macro-specific version of the code. Since completion of our dataset, *AutoMorph* has been further updated to remove scaling errors (see below) and allow for macro-specific 3D processing.

### Data Records

Individual metadata, images, and shape data are provided both in this data report and on the Zenodo data repository (Data Citation 1). The tables in this data report provide relevant metadata including UCMP specimen and locality numbers, locality information, and taxonomy; summary statistics; and information used in technical validation of measurements. Table 1 (available online only) details UCMP specimen numbers and locality numbers, site names, rounded coordinates, and the number of individuals per lot successfully extracted for 2D shape. Each unique site is listed with all associated localities in Table 2 (available online only) along with the number and names of species at that site. Table 3 (available online only) lists the latitudinal range minima and maxima for each species, as well as the UCMP specimen numbers associated with each site and the total number of individuals measured per species. Table 4 (available online only) provides a comparison of *AutoMorph*- and hand-generated measurements. Table 5 (available online only) details the few instances of locality information mismatches present between the UCMP online database and specimen-associated locality information

(see *Usage Notes*). Supplementary Table 1 provides a list of all objects included in the dataset, including an indication of which objects have known locality coordinates.

The data reported on here are all available on Zenodo under a single citation (Kahanamoku *et al.* 2017) and include:

- (1) boxed\_full\_images.zip: 1,256 full-lot (overview) images with boxed object identified by the *segment* module.
- (2) zstacks.zip: image stacks for each of the 1,256 lots processed by *segment*. These image stacks may potentially be used to generate individual 3D height maps.
- (3) 2d\_edfs\_labeled.zip: 12,035 individual EDF images produced by *focus*. A list of the individuals represented in the dataset are provided in Supplementary Table 1 in metadata.zip.
- (4) 2d\_outlines.zip: 12,035 EDF images with 2D outline overlays for quality control.
- (5) 2d\_coordinates.zip: csv files of the outline shape coordinates for the 1,256 lots successfully processed with *AutoMorph*, including one file (all\_coordinates\_limpets.csv) with the coordinates for all 12,035 individuals extracted.
- (6) master\_2d\_measurements.csv: UCMP locality numbers, UCMP lot numbers, locality coordinates, object identification numbers output by *AutoMorph*, and morphometric data for all 12,035 individuals. Each individual has associated full-lot, individual EDF, 2D and outline images, as well as outline shape coordinates, all of which may be found in the zip files listed above. Object identification numbers, which are used to link morphometric measurements with specific individuals, can be found in the column ‘object IDs’. Morphometric data provided includes major axis length, minor axis length, perimeter length, eccentricity, and rugosity, as well as aspect ratio and the height and width of the aspect rectangle. Locality coordinates are rounded to the nearest 0.1 degree to account for potential measurement uncertainty.
- (7) metadata.zip: A directory containing the tables presented in this publication (Tables 1–5 (available online only)) and Supplementary Table 1) in corresponding.xlsx files.

The first data product, boxed\_files.zip, is also available on the UCMP’s photography database (<http://calphotos.berkeley.edu/>). The museum-specific metadata in Table 1 (available online only) can also be found using the UCMP specimen and locality database (<http://ucmpdb.berkeley.edu/>).

## Technical Validation

### Size validation

The lengths and widths of 41 limpets measured by hand were compared to *AutoMorph* outputs of size (Fig. 4, Table 4 (available online only)). *AutoMorph* measurements correspond well to hand measurements of length ( $P < 2.2 \times 10^{16}$ ; intercept  $-0.53$ ) and width ( $P < 2.2 \times 10^{16}$ ; intercept  $-0.56$ ), with an  $R^2$  of 0.999 for both. The largest difference between *AutoMorph* and hand measurements was 7.6 mm, while the mean difference was 0.83 mm. S.d. of the residuals was approximately 0.6 mm for both length and width—well within the variability produced by traditional measurements<sup>24</sup>. We refer interested readers to the Supplementary Text of Hsiang *et al.*<sup>14</sup>, which includes a further comparison between *AutoMorph* and hand measurements for ~100 fossil patellogastropods.

### Image warning labels

*AutoMorph* v2016-02 contained a scaling error that resulted in image scale bars that were consistently 70% larger than true size. The scale bars on all image products from this study are, as a result, 70% too long. Images were appended with warning labels to alert users of the error in scale bar calculation (Fig. 3). The correct pixel size was scaled by a factor of 10 in the 2D image processing steps in order to produce larger output images for visual outline checks. The 2D measurements resulting from this image processing step were subsequently corrected, such that data products 4 and 5 (2D outline coordinates and 2D measurements) are correct as reported.

### Object selection with *segment*

The *AutoMorph* module *segment* produces overview images with each individual object boxed in red (Fig. 3; full set of boxed images available in boxed\_full\_images.zip in data citation). To verify that all objects were being selected for segmenting, the module was first run in ‘sample’ mode, and image selection parameters were visually optimized to allow for proper segmentation.

### Shape extraction with *run2dmorph*

2D extended-depth-of-focus (EDF) images produced by the *focus* module were visually inspected to ensure that the images were properly composited. These EDFs were then passed through the *run2dmorph* module to generate 2D shape and outline extraction. The quality of 2D shape extraction was visually checked for each object using outline-object overlays (see ‘2d\_outlines.zip’ in data citation). Parameters in *run2dmorph* were adjusted as necessary to optimize the fidelity of the 2D outlines extracted.

## Usage Notes

### Expected collection biases

Collection biases are likely to present themselves in lots with small sample sizes, as small lots often introduce collector biases towards larger body sizes (Fig. 5). Users may want to systematically cull the dataset (i.e., perform lot size stepping) in order to account for biases introduced as a result of different collection purposes.

### Curation of locality information

All specimen lots were assigned UCMP locality numbers upon accession into the collection. For three sample lots, the locality information physically associated with the specimens did not match locality information in the online UCMP database. For these sample lots, specimen-associated locality information was considered primary and thus included in the dataset. A record of the mismatched coordinates can be found in Table 5 (available online only).

A number of localities had no associated GPS coordinates, or their locality information was too vague to be associated with GPS coordinates using GEOLocate<sup>16</sup>. Individuals collected from these localities remain in the database but their corresponding locality information was collected directly from specimen-associated catalogue cards and may not be correct.

## References

1. Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M. & West, G. B. Toward a metabolic theory of ecology. *Ecology* **85**, 1771–1789 (2004).
2. Bergmann, C. Über die verhältnisse der wärmeökonomie der thiere zu ihrer grösse. *Göttinger Studien* **3**, 595–708 (1847).
3. Roy, K., Jablonski, D. & Martien, K. K. Invariant size-frequency distributions along a latitudinal gradient in marine bivalves. *Proc. Nat. Acad. Sci* **97**, 13150–13155 (2000).
4. Daufresne, M., Lengfellner, K. & Sommer, U. Global warming benefits the small in aquatic ecosystems. *Proc. Nat. Acad. Sci* **106**, 12788–12793 (2009).
5. Gardner, J. L., Peters, A., Kearney, M. R., Joseph, L. & Heinsohn, R. Declining body size: a third universal response to warming? *Trends Ecol. Evol.* **26**, 285–291 (2011).
6. Berke, S. K., Jablonski, D., Krug, A. Z., Roy, K. & Tomasovych, A. Beyond Bergmann's rule: size-latitude relationships in marine Bivalvia world-wide. *Global Ecol. Biogeog* **22**, 173–183 (2013).
7. Lindberg, D. R. in *Phylogeny and Evolution of the Mollusca* (eds Ponder, W. F. & Lindberg, D. R.) Pp. 271–296 (University of California Press, Berkeley, 2008).
8. Helmuth, B., Mieszzkowska, N., Moore, P. & Hawkins, S. J. Living on the edge of two changing worlds: forecasting the responses of rocky intertidal ecosystems to climate change. *Annu. Rev. Ecol. Evol. Syst.* **37**, 373–404 (2006).
9. Harley, C. D. G. *et al.* The impacts of climate change in coastal marine ecosystems. *Ecol. Lett.* **9**, 228–241 (2006).
10. Hewitt, J. E., Ellis, J. I. & Thrush, S. F. Multiple stressors, nonlinear effects and the implications of climate change impacts on marine coastal ecosystems. *Global Change Biol.* **22**, 2665–2675 (2016).
11. Roy, K., Collins, A. G., Becker, B. J., Begovic, E. & Engle, J. M. Anthropogenic impacts and historical decline in body size of rocky intertidal gastropods in southern California. *Ecol. Lett.* **6**, 205–211 (2003).
12. O'Dea, A., Shaffer, M. L., Doughty, D. R., Wake, T. A. & Rodriguez, F. A. Evidence of size-selective evolution in the fighting conch of prehistoric subsistence harvesting. *Proc. R. Soc. B* **281**, 20140159.
13. Hsiang, A. Y., Elder, L. E. & Hull, P. M. Towards a morphological metric of assemblage dynamics in the fossil record: a test case using planktonic foraminifera. *Phil. Trans. R. Soc. B* **371**, 20150227 (2016).
14. Hsiang, A. Y. *et al.* AutoMorph: Accelerating morphometrics with automated 2D and 3D image processing and shape extraction. *Methods Ecol. Evol.* doi:10.1111/2041-210X.12915 (in press).
15. Ricketts, E. F., Calvin, J., Hedgpeth, J. W. & Phillips, D. W. *Between Pacific Tides*, 5th ed (Sanford Univ. Press, 1985).
16. Rios, N. E. & Bart, H. L. GEOLocate (Version 3.22). *Tulane Museum of Natural History* (Belle Chasse, LA, 2010).
17. Merriam, J. C. An outline of progress in paleontological research on the Pacific Coast. *Univ. Calif. Publ., Bull. Dept. Geol* **12**, 237–266 (1921).
18. Coan, E. *James Graham Cooper, Pioneer Western Naturalist* (Univ. Press, 1982).
19. Coan, E. & Roth, B. The malacological taxa of Henry Hemphill. *Veliger* **29**, 322–339 (1987).
20. Dall, W. H. Professor Josiah Keep [Obituary]. *Science* **34**, 371 (1911).
21. University of California Museum of Paleontology. *UC Davis Locality Ledgers, Coan Collection*, Vol. 1-3 (UCMP, Berkeley, CA, 1960–1970).
22. Test, A. R. G. Ecology of California Acmaea. *Ecology* **26**, 395–405 (1945).
23. Fritchman, H. K. A study of the reproductive cycle in the California Acmaeidae (Gastropoda). *Veliger* **4**, 41–47 (1961).
24. Lindberg, D. R. in *Proceedings of the Michigan Morphometrics Workshop*, (eds Rohlf, F. J. & Bookstein, F. L.) Pp. 301–301 (Univ. Mich. Mus. Zool. Spec. Pub. 2, 1990).

## Data Citation

1. Kahanamoku, S. S. *et al.* Zenodo <http://doi.org/10.5281/zenodo.801765> (2017).

## Acknowledgements

We would like to thank Kaylea Nelson and Brian Dobbins at Yale for their invaluable technical expertise. This work was supported by an American Chemical Society PRF Grant #55837-DNI8 to P.M.H. and A.Y.H., as well as Yale University internal support to P.M.H., A.Y.H., and S.S.K.

## Author Contributions

S.S.K., P.M.H., and S.F. designed the study. S.S.K. collected image data. S.S.K. and A.Y.H. processed images and compiled all data products, and P.M.H. provided a final quality control on all data products. E.C.C. curated all samples. All authors contributed to the text and data processing and validation.

### Additional Information

Tables 1–5 are only available in the online version of this paper.

**Supplementary Information** accompanies this paper at <http://www.nature.com/sdata>

**Competing interests:** The authors declare no competing financial interests.

**How to cite this article:** Kahanamoku, S. S. *et al.* Twelve thousand recent patellogastropods from a northeastern Pacific latitudinal gradient. *Sci. Data* 5:170197 doi: 10.1038/sdata.2017.197 (2018).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

The Creative Commons Public Domain Dedication waiver <http://creativecommons.org/publicdomain/zero/1.0/> applies to the metadata files made available in this article.

© The Author(s) 2018